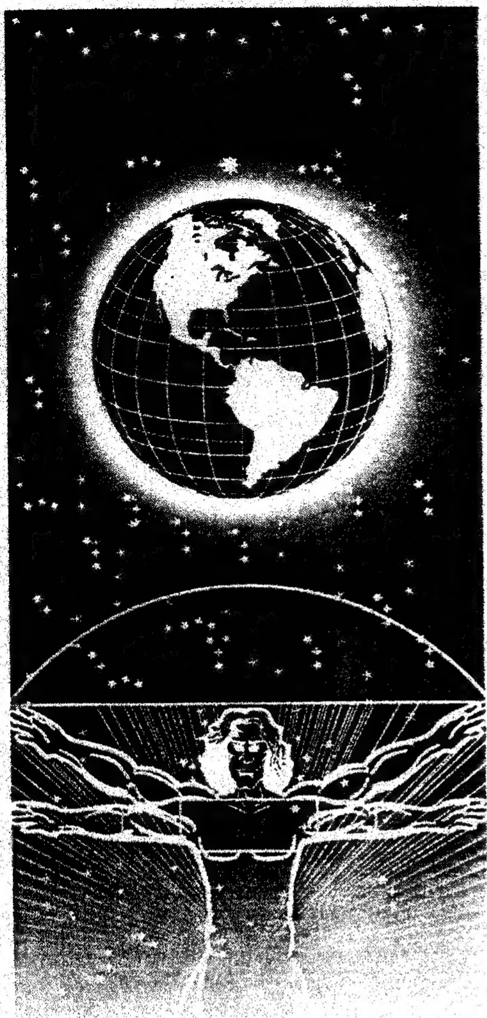
AFRL-HE-AZ-TR-1998-0086



UNITED STATES AIR FORCE RESEARCH LABORATORY

A COGNITIVE CORRELATES ANALYSIS OF SITUATION AWARENESS

William C. Tirre

Air Force Research Laboratory
Warfighter Training Research Division
Aircrew Performance Branch
7909 Lindbergh Drive
Brooks Air Force Base TX 78235-5352

Leo J. Gugerty

GALAXY SCIENTIFIC CORPORATION
7334 Blanco Road
San Antonio TX 78216

April 1999

AIR FORCE MATERIEL COMMAND
AIR FORCE RESEARCH LABORATORY
HUMAN EFFECTIVENESS DIRECTORATE
WARFIGHTER TRAINING RESEARCH DIVISION
6030 South Kent
Mesa AZ 85212-6004

NOTICES

Using Government drawings, specifications, or other data included in this document for any purpose other than Government-related procurement does not in any way obligate the US Government. The fact that the Government formulated or supplied the drawings, specifications, or other data, does not license the holder or any other person or corporation, or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

The Office of Public Affairs has reviewed this report, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This paper has been reviewed and is approved for publication.

WILLIAM C. TIRRE
Project Scientist

DEE H. ANDREWS
Technical Director

JERALD L. STRAW, Colonel, USAF
Chief, Warfighter Training Research Division

Please notify AFRL/HEOA, 2509 Kennedy Drive, Bldg 125, Brooks AFB, TX 78235-5118, if your address changes, or if you no longer want to receive our technical reports. You may write or call the STINFO Office at DSN 240-3877 or commercial (210) 536-3877, or e-mail Shirley.Walker@afrlars.brooks.af.mil

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)

2. REPORT DATE
April 1999

3. REPORT TYPE AND DATES COVERED
Final - October 1995 to August 1998

4. TITLE AND SUBTITLE

A Cognitive Correlates Analysis of Situation Awareness

5. FUNDING NUMBERS

F41624-95-C-5007
PE - 61102F
PR - 2313
TA - T1
WU - 47

6. AUTHOR(S)

William C. Tirre
*Leo J. Gugerty

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Air Force Research Laboratory
Human Effectiveness Directorate
Warfighter Training Research Division,Aircrew Performance Branch
7909 Lindbergh Drive, Brooks Air Force Base, TX 78235-5352
*Galaxy Scientific Corporation, 7334 Blanco Road, San Antonio TX 78216

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Air Force Research Laboratory
Human Effectiveness Directorate
Warfighter Training Research Division
6030 South Kent
Mesa AZ 85212-6004

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

AFRL-HE-AZ-TR-1998-0086

11. SUPPLEMENTARY NOTES

Air Force Research Laboratory Technical Monitor: Dr William C. Tirre, (210) 536-2027

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

One of the most important abilities needed to perform real-time tasks such as piloting an aircraft or driving a car involves monitoring and comprehending the rapidly changing situations in these tasks, that is, maintaining situation awareness (SA). In this research a model of situation awareness in driving was used to derive a set of individual differences measures. These individual differences measures were embedded in a highway driving simulation (PC-DriveSim) that presented scenarios varying in the type of traffic hazard encountered. The SA model was also used to derive a set of potential cognitive correlates that could be measured and used to predict SA levels in individual subjects. Factor analysis of the SA measures indicated that a single factor provided adequate fit to the data. Then in a set of three correlational studies, a set of cognitive correlates or predictors of situation awareness was identified and refined. In the third study, six predictors were related to an SA composite score resulting in an equation that accounted for 50% of the variance. The variables that made significant unique contributions to the equation were time-sharing, temporal processing (rate extrapolation and duration discrimination), visual search, and dynamic visual processing.

14. SUBJECT TERMS
Aptitude-treatment interaction; intelligent tutoring personality; Learning; Learning aptitudes; Motivation; Real-time tasks; Simulation; Situation awareness; SA; Task performance

15. NUMBER OF PAGES
31

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev 2-89)
Prescribed by ANSI Std Z-39-18

A COGNITIVE CORRELATES ANALYSIS OF SITUATION AWARENESS

INTRODUCTION

One of the most important abilities needed to perform real-time tasks involves monitoring and comprehending the rapidly changing situations in these tasks, that is, maintaining situation awareness (SA). People differ in their ability to maintain SA. In a number of experiments, we measured these individual differences in SA abilities and also measured participants' abilities at a variety of basic cognitive tasks. Discovering the basic cognitive abilities that correlate with SA abilities gives us information about the cognitive processes used in maintaining SA. Thus, this research can help us to refine our understanding of the global construct of SA. Also, identifying tests of the cognitive correlates of SA can help in selecting individuals who are likely to perform well in real-time tasks.

We define a real-time task as one where: (1) the external task environment changes continuously, with some changes being beyond the operator's control, and (2) the operator must allocate attention among multiple subtasks (time-sharing or multitasking). SA is the activated knowledge used to perform a real-time task, i.e., the knowledge in working memory or easily available to working memory (cf. Ballas, Heitmeyer & Perez, 1992). Situation assessment refers to the cognitive processes used to maintain SA. These processes involve comprehending a dynamic, multifaceted task situation.

One of the major models of real-time task performance, Klein's recognition-primed decision model, emphasizes the importance of SA (Kaempf, Klein, Thordsen & Wolf, 1996; Klein, 1993). This model suggests that real-time operators do very little problem solving. That is, once they have recognized (or comprehended) a situation, they consider very few response alternatives. Their responses are usually determined by simple condition-action rules (Orasanu, 1996; Orasanu & Fischer, 1997). Thus, the quality of operators' responses will depend largely on how well they have assessed the situation.

In support of this claim, naturalistic studies have demonstrated that the most frequent cause of errors in real-time tasks is errors in SA (Endsley, 1995a; Hartel, Smith, & Prince, 1991). For example, in a study of 420 automobile accidents, Shinar (1993) found that the most frequent cause was what the author called "recognition" errors such as improper lookout, inattention, and internal distraction, with errors in response selection and execution being less frequent.

We have studied SA using the real-time task of driving. Figure 1 presents our analysis of some of the cognitive processes involved in situation assessment during driving, and includes these processes in a simple framework for driving decision making. Drivers must maintain knowledge of route location needed for navigation, knowledge of nearby traffic needed for maneuvering (local scene perception), knowledge of spatial orientation (e.g., lane position) needed for path tracking, and knowledge of the status of their vehicle. For one of these driving

processes, local scene comprehension, we have indicated a sub-process, scene perception and projection, which refers to the driver's ability to perceive the locations and speeds of nearby vehicles and project these into the future. In contrast to scene perception and projection, local scene comprehension focuses on a driver's ability to understand the meaning inherent in the nearby traffic, e.g., to identify potential hazards. Subprocesses for other driving processes, such as maintaining navigation knowledge, have not been indicated in Figure 1 because our analysis has not yet focused on these processes.

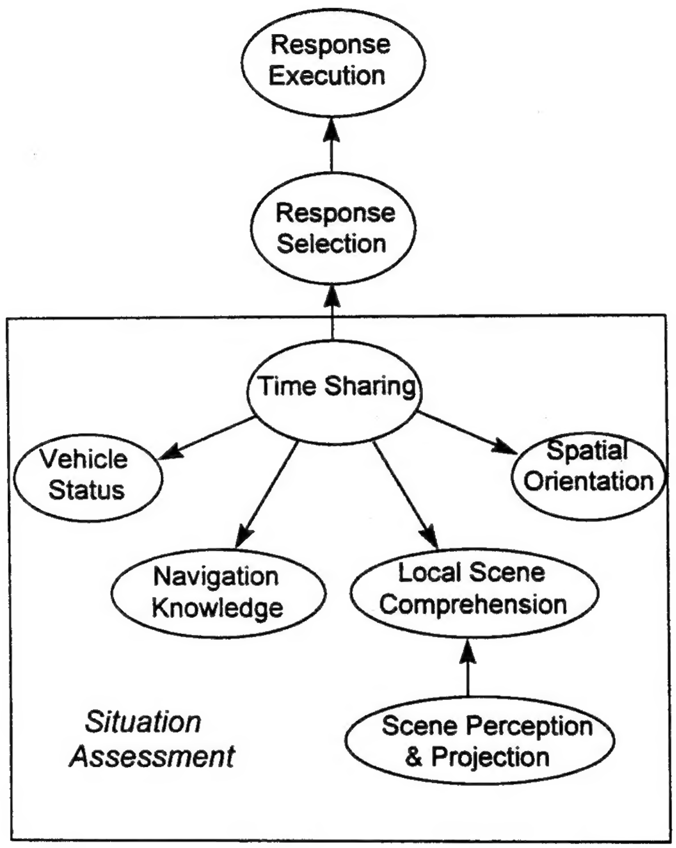


Figure 1. Cognitive Processes Used in Driving

The time-sharing process in Figure 1 refers to the driver's ability to allocate attention among the four situation-assessment processes. Finally, the situation assessment processes are shown in the context of later stages of driving decision making, response selection, and execution.

Our investigation of individual differences in SA abilities has focused on local scene comprehension, scene perception and projection, and time-sharing. In addition, we have also measured abilities at a variety of basic cognitive tasks to see which of these abilities predicted SA abilities. The cognitive abilities we measured included working memory, temporal processing, visual processing, and perceptual-motor coordination. The reasoning behind selection of these predictors is as follows:

- Working memory has a limited capacity that must be shared between current operations and temporary storage of intermediate results and freshly encoded data. It is the central bottleneck in information processing and its capacity varies considerably across individuals. Thus, we should expect SA to be limited by working memory capacity.

- Visual processing is likely an important part of SA for driving in that safe drivers scan the environment, checking mirrors and forward view to monitor the locations and actions of other vehicles.

- Temporal processing is probably a component of performance in dynamic visual environments such as driving. For example, temporal processing is probably involved in estimating velocities, in maintaining a safe stopping distance between cars, and in noticing and passing slower vehicles on the highway.

- Lastly, perceptual-motor coordination might be involved in SA, not because maintaining SA requires motor control, but because it requires time-sharing ability. In the SA tests used in this study, participants monitored multiple vehicles so that they could later perform a number of tasks: identifying hazardous vehicles, recalling vehicle locations, and making driving actions that avoid hazards. Monitoring multiple vehicles in order to perform multiple tasks required time-sharing. Similarly, the perceptual-motor coordination tasks we used required time-sharing in that the participants had to view moving stimuli and coordinate movements of hand and feet.

Previous Research on Individual Differences in Situation Awareness

Previous research has examined the cognitive correlates of SA in aircraft pilots.[1] For example, Carretta, Perry, and Ree (1996) reported a study in which a large battery of cognitive, perceptual, and perceptual-motor tasks were administered to 171 F-15 pilots. The dependent variable was the first unrotated principal component found on a set of peer and supervisor ratings of SA.[2] The dependent variable was predicted well by F-15 flying hours (experience) but the only individual difference variable that had any incremental validity was a general cognitive ability (*g*) composite.[3] When F-15 flying experience was partialled out of both the dependent

[1] Our literature search on individual differences in SA identified only three publications in the scientific literature, all of which pertained to SA in aircraft pilots. Our working assumption is that there is considerable commonality to the cognitive and perceptual processes that support SA in a variety of operator tasks. Thus, we believe that what we learn about SA in piloting aircraft will be relevant to SA in driving automobiles, and vice versa.

[2] Someone might question whether the Carretta et al. dependent variable actually reflected SA. One might doubt it is meaningful to ask a person to rate the quality of another person's unobservable mental events. An alternative interpretation of their dependent variable was that it reflected general airmanship (piloting skill) as perceived by other pilots who presumably had ample occasion to make informed judgments. One bit of data that might be used to support the SA interpretation of the rating scale is found in Bell and Waag (1995) who found a correlation of .6 (N = 40) between the SA rating scale and ratings of SA performance in simulated fighter missions.

[3] Spearman (1904) presented the first formal theory of general cognitive ability or general intelligence. He argued that performance on ability tests was a function of (1) a general intellectual factor which underlied all cognitive performance to some extent, and (2) a specific factor unique to the task. Interestingly, Spearman (1923) anticipated modern cognitive psychology when he speculated that general intelligence reflected individual differences in "mental energy" or attention.

variable and the predictor variables, significant correlations were found for working memory, divided attention tests, and two perceptual-motor tests.

Objective, performance-based measurement of SA, perhaps in a flight simulator, would probably provide the basis for a deeper understanding of SA. Along these lines, Endsley and Bolstad (1994) examined correlates of SA using the Situation Awareness Global Assessment Technique (SAGAT) and a battery of 18 cognitive, perceptual, and perceptual-motor tests. Because of the small sample (N = 21), strong conclusions about correlations with SAGAT are not warranted. But it is interesting to note that a perceptual-motor tracking task correlated 0.72 with SAGAT. Endsley and Bolstad suggested that pilots with superior perceptual-motor abilities had spare attentional capacity that could be devoted to situation assessment; consequently they demonstrated better SA.

Other research has investigated SA in the context of artificial tasks that vary in the degree to which they resemble real-world tasks. In this literature, SA is often described as a high-level attention management ability that is distinct from more elementary cognitive processes (see for example, Hopkin, 1993). Along this line, O'Hare (1997) reported a study in which a small sample of adult males (N = 24) was administered the WOMBAT Situational Awareness and Stress Tolerance Test (Roscoe, 1993; Roscoe & Corl, 1987) and tests from the Walter Reed Performance Assessment Battery which corresponded to component tasks of the WOMBAT. The WOMBAT is a complex time-sharing task that requires the examinee to divide attention among multiple tasks on separate screens. To score well on the WOMBAT, examinees must be able to quickly assess which task has priority at any given moment and direct attention to it. O'Hare found that only one subtest, pattern recognition[4], was consistently correlated with WOMBAT performance through the 60-minute duration of the task. During the first 10 minutes of practice, the correlation was .59, and during the last 10 minutes, .57 ($p < .01$). O'Hare presented analyses that suggested WOMBAT performance became less dependent on computer game experience and elementary cognitive processing abilities with practice.

Our reanalysis of the data did not support this conclusion: in the final 10 minutes of task performance, both computer experience and pattern recognition scores continued to make unique contributions to R^2. Given the small sample size, and the small differences between predictor-criterion correlations for the first and final 10 minutes of practice, the prudent strategy would be to use the final 10 minutes of WOMBAT performance as the dependent variable. If one considers the .63 correlation between the first 10 and final 10 minutes of WOMBAT performance to be a reliability estimate, the disattenuated correlation between pattern recognition and WOMBAT performance is .72. If anything, this correlation indicates that SA (if we accept WOMBAT as a measure of SA) is related to visual recognition memory even after 60 minutes of practice at the SA task.

[4] The pattern recognition test presented a random arrangement of 16 asterisks for 1.5 seconds followed by a retention interval of 3.5 seconds. Then a second pattern is presented which has two randomly chosen asterisks changed in position. The examinee must decide if the study and test stimuli are the same or different.

In complex real-time environments, correlation of task performance with cognitive ability might increase rather than decrease with practice. Rabbit, Banerji, and Szymanski (1989) reported a study in which 56 males were administered five 1-hour training sessions on the Space Fortress task over five days. They were also administered a standard intelligence (general cognitive ability) test. The Space Fortress task is a complex video game that involves manipulating a spacecraft to attack a space fortress that is trying to defend itself. Perceptual-motor as well as purely cognitive abilities are required to perform the task. Performance improves slowly over hours of practice, but participants generally find it engaging. Space Fortress was designed to develop general workload-coping and attention-management skills and there is some empirical validation that training transfers to real aircraft piloting (Gopher, Weil, & Bareket, 1992; Hart & Battiste, 1992). Rabbit et al. found that Space Fortress performance correlated .28 with general cognitive ability in the first hour of practice, and .69 in the final or fifth hour of practice. The interpretation Rabbit et al. (1989, p. 254) gave these data clearly links general cognitive ability to SA. High correlations between general cognitive ability test scores and game performance "... may occur because people who can master most of the wide range of problems included in IQ tests can also more rapidly learn to master complex systems of rules, to attend selectively to the critical portions of complex scenarios, to make rapid and correct predictions of imminent events and to prioritize and update information in working memory."

Some critics might unfairly dismiss results with Space Fortress because the task is unrealistic, but ongoing research by Tirre, which made use of a more realistic aircrew task, has yielded similar results. Tirre investigated the cognitive correlates of a synthetic task designed to simulate the B-1 defensive systems operator task (the DSO Analog task), but which did not require specialized knowledge. DSO Analog involved identifying threats (called enemies in the game) and selecting the appropriate defense. The subject's ship moved forward at a constant rate but subjects were able to move their ships left or right. When an enemy became within range of the subject's ship, the enemy was able to shoot missiles and/or lasers at the ship depending on the type of enemy. Some enemies could also launch more enemies. If the subject applied the correct defense while an enemy was within range, the enemy could not attack the ship and would harmlessly pass by. The task required comprehending a complex set of rules, noticing when the automation misidentified enemies, and selecting appropriate actions in real time. As with Space Fortress, scores started out very low (even negative) and improved with hours of practice.

Tirre found that correlations of DSO performance with a working memory capacity composite score steadily climbed from .38 for the first hour to .53 for the fourth and final hour (N = 130). Likewise, correlations with a general cognitive ability score derived from paper-and-pencil tests climbed from .31 to .53. The reason for the increase in correlation between task performance and the ability measures is likely due to a general increase in DSO internal consistency. Internal consistency reliability increased from .76 to .92 with practice, which probably reflects a stabilization of strategies adopted by the examinees to perform the DSO task.

Joslyn and Hunt (1998) investigated the cognitive correlates of three realistic tasks that seemed to require SA in that they required operators to monitor and classify a number of changing situations and allocate scarce resources to the situations based on their classifications. These tasks simulated aspects of the job of a public safety (911) dispatcher, a public safety call

receiver, and an air traffic controller. In four studies, the performance of both college students and professional operators (911 dispatchers) on these realistic criterion tasks were predicted, with correlations ranging from .50 to .70, by performance on the abstract decision-making (ADM) task. ADM, also developed by Joslyn and Hunt (1998), was a dynamic, content-free task that required monitoring and classifying multiple events that overlapped in time. ADM probably measures what we have termed time-sharing ability, in that it required participants to prioritize and switch attention among multiple overlapping events. It also seems to measure working memory ability, in that participants must remember the rules for classifying events; the events that still need to be classified; and what features of those events have been identified. In a fifth study, Joslyn and Hunt found that ADM predicted both initial and well-practiced performance on the 911 dispatch task.

To summarize, previous research on the cognitive correlates of SA has measured SA ability using questionnaire data (Carretta et al., 1996) and from performance on moderate and high fidelity simulators (Endsley & Bolstad, 1994; Joslyn & Hunt, 1998) and synthetic laboratory tasks (O'Hare, 1997; Rabbit, Banerji & Szymanski, 1989). Although some of these studies used too few participants to warrant strong conclusions, the general finding is that SA is related to working memory capacity, general intelligence (g), time sharing ability, perceptual-motor ability, and visual recognition ability.

We next describe the SA tests we used and then present the results of three studies of the cognitive correlates of SA. Our studies were similar to the Endsley and Bolstad study in that SA was measured objectively. They were dissimilar in that SA was measured in a less complex but still dynamic environment (simulated highway driving) on a sample of adults with a broader range of abilities than considered in either the Carretta et al. (1996) or Endsley and Bolstad (1994) studies.

Descriptions of Situation Awareness Tests

Our SA tests focused on the driving subtasks of local scene comprehension and scene perception and projection shown in Figure 1. Each test required participants to monitor the movements of the nearby vehicles in a simulated driving task and to make judgments and actions based on these movements.

The driving task was performed on a PC-based driving simulator. The simulator showed 3-dimensional animated driving scenes in a window that filled a 17-inch computer screen. The participant saw the front view from the driver's perspective and also the rearview, left-sideview, and right-sideview mirrors (see Figure 2). All scenes showed traffic on a three-lane divided highway, with all cars moving in the same direction. Participants watched animated scenes lasting from 18 to 35 seconds and were instructed to imagine that their simulated car was on autopilot. At the end of each scene, participants' knowledge of the traffic vehicles was probed using a number of methods.

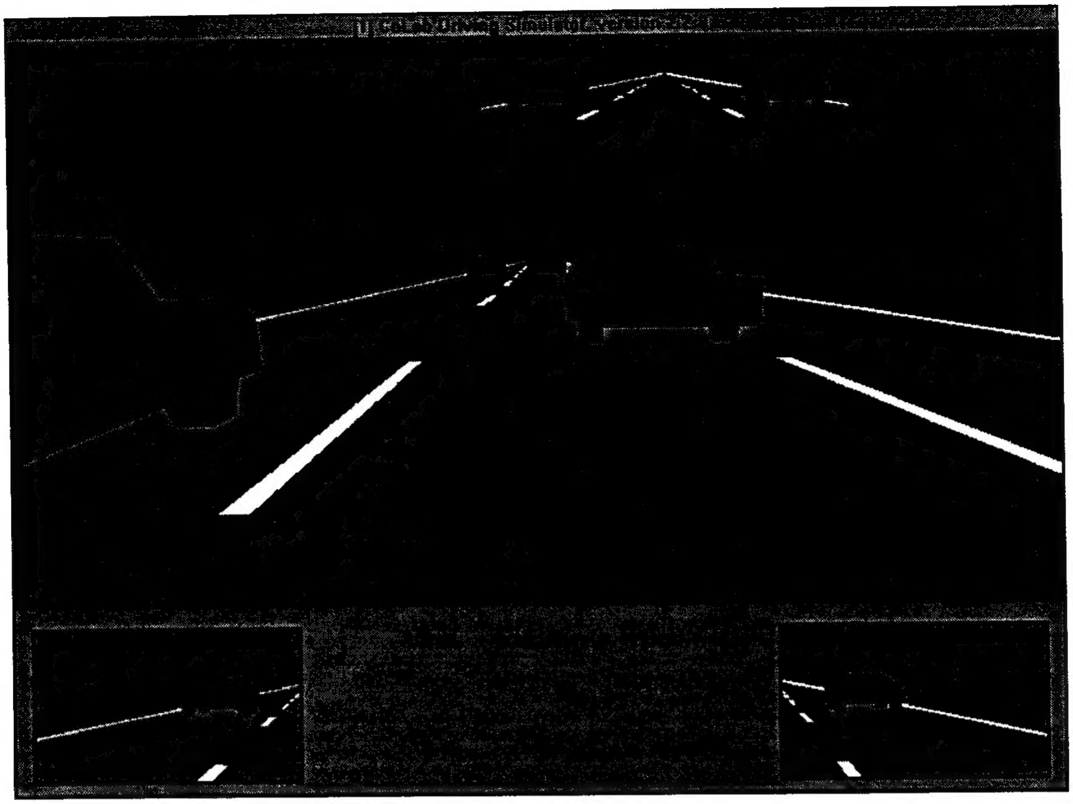


Figure 2. Three-dimensional scene from driving simulator. The actual scenes were in color.

In the recall probes, the moving scene disappeared and participants, using the mouse, indicated the locations of the traffic cars at the end of the scene on a bird's-eye view of the road. The bird's-eye view showed the road 20 car lengths ahead of the driver and 10 lengths behind, and the driver's car (in the correct lane). After participants finished recalling the car locations for a scene, they received feedback indicating the correct final car locations for that scene.

In the time-to-passage probes, the driver's car was in the center lane and was either overtaking or being overtaken by two cars in the right and left lanes. In either case, the moving scene stopped before the two traffic cars reached the driver. The participant then judged which of the two traffic cars would reach his or her car first.

In the performance probes, participants could make driving responses while viewing the moving scenes; that is, they could override the autopilot. On some trials, an incident would occur that required a driving response, for example, a car would move into the driver's lane ahead of the driver while moving slowly enough that it would hit the driver. Other hazards approached the driver from the rear. Participants could make four responses to avoid hazards: accelerate, decelerate, move to the left lane, or move to the right lane. They indicated these responses with the up, down, left, and right arrow keys, respectively. When participants pressed an arrow key during a hazard, the scene ended and feedback was displayed concerning the correctness of the response. Correct performance required the participant to avoid hazard cars without hitting any other traffic cars and to refrain from responding on catch (non-hazard) trials.

In the scene-interpretation probes, no driving (arrow-key) responses were required. At the end of the moving scene, the participants saw a bird's-eye view of the road that included the correct locations of the traffic cars and were asked a question that required them to identify potentially hazardous cars. The questions included: Which car was driving erratically? Which car was tailgating? Which car is most likely to change lanes now? Which car is driving fastest (or slowest) now? Which two cars are on a collision course now?" Participants clicked on the car or cars that answered the question and then received feedback.

The time-to-passage probes were always done in a separate block of trials. However, the other three probes were usually done in the same block. A participant would watch a moving scene and make a driving response (performance probe) if necessary. Then, at the end of the scene, the participant might be asked to recall car locations or identify hazards. Thus, on the blocks with three types of probes, participants had to monitor the information required for all three probes on each trial. This required time-sharing ability.

We hypothesized that the recall probes and time-to-passage probes would assess participants' abilities to perceive speeds and distances and to project speed and distance information into the future to predict future vehicle locations. These are the abilities referred to as scene perception and projection in Figure 1. The scene-interpretation probes and performance probes required participants to do more than perceive and project speeds and distances. Participants also had to comprehend the meaning of the speed and distance information to identify and avoid hazards. Thus the latter two probes were hypothesized to reflect local scene comprehension abilities.

The specific ability measures used for each probe type were as follows. To evaluate the "goodness" of participants' recall data, first, a computer algorithm matched the cars recalled by participants with the actual locations of cars at the end of each trial, and also identified nonrecalled cars and false alarms. Once the matching was done for a trial, the composite recall error was calculated based on the average distance between recalled cars and the actual cars they were matched to and on the number of nonrecalled and false-alarm cars. (See Gugerty (1997) for details concerning how this and other SA measures were calculated.)

The time-to-passage probe data were scored in terms of the percentage of trials on which participants answered correctly. The scene-interpretation data were scored in the same manner.

Two SA measures were derived from the performance-probe data. The first was hazard detection. This was calculated using the A' nonparametric, signal-detection measure of sensitivity (Grier, 1971). On each signal (hazard) trial, a response interval was defined as beginning when a car entered the driver's lane on a trajectory that would hit the driver and ending when it was too late for the driver to avoid the oncoming car. Following the procedure of Watson and Nichols (1976) for measuring sensitivity with continuous signal-detection tasks, we defined catch-trial response intervals that were equal in duration to those on hazard trials. A hit was defined as any arrow-key response, even an incorrect response, during the response interval of a hazard trial. A false alarm was any arrow-key response during the response interval of a catch trial. For all trials, responses before the response interval, which were infrequent, were ignored in this analysis.

We attempted to construct the hazard detection measure so that it assessed SA ability and not decision making or response execution ability, even though any performance-based measure is affected by all stages of the perception-action cycle. When participants responded incorrectly to a hazard car, this showed that they were aware of the hazardous situation, but selected and executed an inappropriate avoidance response. Therefore, by defining even incorrect responses to hazards as hits in this measure, we hoped that it would reflect participants' ability to detect hazards (an aspect of SA) more than their decision-action abilities.

The hazard detection measure focuses on participants' awareness of vehicles in front of and behind their car, because the hazardous cars always entered the driver's lane from a side lane and then approached the driver. The second SA measure derived from the performance-probe data, <u>blocking-car detection</u>, focused on participants' awareness of blocking cars to their immediate right and left. These cars were usually completely within the participants' blind spot. Participants could usually only know about blocking cars by remembering that a car had entered the blind spot and had not left it. On a trial where the hazard car approached from the front and there were blocking cars to the right and left, participants were considered as detecting one of two blocking cars if they went right or left. Overall, blocking-car detection was estimated by the ratio of the total number of blocking cars avoided over the total number of blocking cars.

As in the hazard detection measure, scoring high on the blocking-car detection measure does not depend on making a correct response in terms of global task performance. In the above example, participants would be credited with 50% blocking-car detection on a trial where they crashed. Thus, blocking-car detection should reflect participants' awareness of blocking cars more than their decision-action processes.

EXPERIMENT 1

This study was conducted in conjunction with a larger factor analytic study of the Cognitive Abilities Measurement (CAM) battery (Kyllonen, 1994). CAM attempts to comprehensively measure the human abilities essential to the acquisition of knowledge and skill.

Method

Participants. Participants were hired from temporary employment agencies. In the larger study, 230 participants of both sexes between the ages 18 to 30 were administered the CAM and the Armed Services Vocational Aptitude Battery (ASVAB) over five days. A subset (N = 34) was also administered the driving simulator task.

Abilities Tests. In the driving simulator session, participants completed 84 trials with both performance and recall probes and 84 trials with only recall probes. CAM, a battery of 59 computer-administered tests, was created through the use of a taxonomy. The six rows of the taxonomy reflect the major abilities suggested by cognitive psychology, *viz.*, working memory, processing speed, induction, declarative knowledge, declarative (associative) learning, and procedural (skill) learning. The columns reflect three information domains suggested by psychometric analyses, *viz.*, verbal, quantitative, and spatial. Three or four tests were given for

each cell of the taxonomy. The ASVAB is the entrance examination for the US armed services. It is a paper-and-pencil test that includes 10 subtests.

Procedure. Participants were tested in groups of about 20 who completed the tests over a week-long period.

Results and Discussion

As discussed in Gugerty (1997, Experiment 1), participants' scores for the driving simulator tests were reliable and above chance. Test battery scores for the cognitive abilities tests were reduced to a manageable number in two steps. First, for the 230 participants in the full study, 18 cell scores for the six rows and three columns of the CAM taxonomy were formed, e.g., all spatial working memory tests were combined into one score. Second, the 18 CAM and 10 ASVAB scores were factor analyzed in separate runs. Factor analysis resulted in three CAM factors -- g/working memory (*g/WM*), processing speed, and declarative knowledge -- and two ASVAB factors -- *g* and clerical speed. Each factor set was orthogonally rotated using the quartimax procedure, which emphasizes the amount of variance explained by the first factor. The first CAM factor, *g/WM*, had its highest loadings on the working memory and procedural learning tasks. This finding is consistent with prior research (e.g., Kyllonen & Christal, 1990; Tirre & Pena, 1993) that indicates working memory and tasks with heavy working memory requirements might be the core of the psychometric phenomenon known as *g*. The first ASVAB factor had its highest loadings on tests requiring quantitative reasoning skills, *viz*., arithmetic reasoning and math knowledge, both of which benefit from education. As such, ASVAB *g* reflects crystallized instead of fluid ability (Cattell, 1971).

The SA measures were highly correlated with the CAM *g/WM* factor and slightly less so with the ASVAB *g* factor. Hazard detection was correlated .60 ($p < .001$) with *g/WM*, and .39 with ASVAB *g* ($p < .03$). Likewise, blocking-car detection correlated .74 ($p < .001$) with CAM *g/WM* and .76 ($p < .001$) with ASVAB g. Composite recall error correlated -.73 with *g/WM* ($p < .001$) and -.50 with ASVAB *g* ($p < .003$). None of the remaining ability factors correlated significantly with the SA measures. These results are consistent with the hypothesis that SA depends critically on the working memory system (Endsley, 1995b).

EXPERIMENT 2

The second experiment was conducted as part of a larger factor analytic study of cognitive, perceptual-motor, and temporal processing abilities. Each participant completed a session in the driving simulator. In addition, computer-based batteries of cognitive, perceptual-motor and temporal tests, and one paper-and-pencil battery (the Air Force Officer Qualifying Test, or AFOQT) were administered.

Method

Participants. Participants were hired from temporary employment agencies. The 88 participants included 61 males and 27 females ranging in age from 18 to 30 years.

Abilities Tests. In the driving simulator session, participants completed 148 trials with only performance probes. In this experiment, we were interested in whether the frequency of hazards in the performance probes affected participants' ability to detect hazards. There were two groups: one that experienced hazards on 75% of the trials and a second that experienced hazards on only 25% of the trials. Both groups first experienced practice in which hazards occurred on 50% of the trials. As it turns out, hazard frequency did not have a significant effect on any of the dependent variables (Gugerty & Tirre, 1996), so we collapsed across groups for the correlational analyses.

The computer-administered cognitive battery was a subset of the CAM 4.1 battery consisting of tests of working memory, processing speed, induction, declarative knowledge, declarative learning, and procedural learning (Kyllonen, 1994). We used only the working memory tests in our analysis. These consisted of a spatial test where participants memorized, combined, and then recalled stick figures; a quantitative test where participants studied lists of numbers, made computations with them, and then recalled them; and a verbal test where participants studied a series of short sentences and then created word lists based on the relations in the sentences.

The perceptual-motor battery consisted of 17 tests designed to measure four of the Fleishman factors: multilimb coordination, control precision, rate control, and response orientation (Fleishman & Quaintance, 1984). In this experiment we used the multilimb coordination tests. In the "center-the-ball" task, the participant adjusted the horizontal motion of a drifting circle with foot pedals and the vertical motion with the joystick so as to keep the circle over a target. In "pop the balloons," the participant used the foot pedals and joystick to move a sight over a moving target circle and then "popped" the balloon with a trigger press. In the Mashburn task, participants used the joystick and foot pedals to move and maintain three cursors so that each was at a separate target point. Each trial was completed when a two-second period had elapsed in which all three cursors were at their target points. In other tasks, participants used both hands (two joysticks) to track a moving object and to move an object along a path.

In the temporal battery, participants estimated when a growing line or an ascending digital clock would reach a target, selected which of two growing lines or ascending clocks would reach a target first, and estimated when short time intervals had ended.

Visual search ability was estimated based on three subtests of the AFOQT: table reading, block counting, and scale reading. Each of these subtests appears to require the participant to visually scan a printed stimulus (a table of numbers, a picture of stacked blocks, or a scale with markings) for a certain item or set of data. Each subtest requires attention to visual detail and in a way might be regarded as an oculomotor test since eye movements must be carefully controlled.

Procedure. Participants were tested in groups of about 20 who completed the test batteries over a week-long period. Each battery was given on a fixed day or days. The order of tests within a battery was randomized for each participant.

Results and Discussion

Because the participant-to-variable ratio was too small to warrant factor analysis of the total dataset, we created composite scores corresponding to factors found in previous analyses of the CAM 4.0, perceptual-motor, temporal, and AFOQT batteries in which sample size was no problem. Considering CAM first, we created a composite for working memory, corresponding to a factor found by Kyllonen (1993). For the perceptual-motor battery we simply created a multilimb coordination composite corresponding to Fleishman and Quaintance's (1984) factors since we had deliberately attempted to simulate original apparatus tests used by Fleishman in his factor analytic research. We made a separate temporal processing composite, since Tirre (1997) found that a similar factor loaded by two of the temporal processing tests was distinct from four perceptual-motor factors including multilimb coordination. Factor analyses by Goff and Tirre (1995) found that AFOQT variance resolved down to seven factors, including visual search.[5] Based on these findings, the visual search composite was created from the table reading, scale reading, and block counting subtests of the AFOQT.

The SA measures, hazard detection, and blocking-car detection were significantly intercorrelated ($\underline{r} = .42$, $\underline{p} < .001$). The correlations of the four ability predictors with the SA dependent variables (see Table 1) indicate that: (a) hazard detection is correlated significantly with all predictors, with working memory having the highest correlation, and (b) blind spot avoidance is correlated significantly with three of the four predictors. However, substantial intercorrelations of the predictor variables and the small sample size suggest caution in interpreting these correlational patterns.

Multiple regression analyses with the dependent variables reflect the pattern displayed in Table 1. That is, working memory was the only significant predictor ($\underline{p} < .05$) in equations for hazard detection ($\underline{R} = .46$), and both working memory and multilimb coordination were significant predictors of blocking-car detection ($\underline{R} = .51$). These two predictor variables were correlated .47, and so we might expect a substantial common contribution from them in a regression equation. In a two-variable equation predicting hazard detection, the unique contributions of working memory and multilimb coordination to the explained variance were .11 and .01, respectively; while the common contribution was .08. In the equation predicting blocking car detection the corresponding values were .11, .03, and .10.

In summary, the findings suggest that SA ability depends primarily on working memory and multilimb coordination ability, and to a lesser extent on temporal processing and visual search ability. As mentioned above, multilimb coordination may predict the non-motor SA tasks because it taps into participants' time-sharing ability.

[5] Carretta and Ree (1996) reported a six-factor solution that included a "perceptual speed" factor. The Goff and Tirre "visual search" factor was identical to the Carretta and Ree perceptual speed factor. Goff and Tirre preferred the visual search label because perceptual speed misleadingly implies near visual threshold processing

Table 1. Correlations Between Predictor and Situation-Awareness Measures for Experiment 2

Predictor	Situation Awareness Measure: Hazard detection	Situation Awareness Measure: Blocking-car detection
Working memory	.44***	.47***
Multilimb coordination	.31**	.37***
Visual search	.22*	.21
Temporal processing	.35***	.31**

Note. N = 88. *p < .05. **p < .01. ***p < .001.

However, there are alternative interpretations of the Experiment 2 findings. The multilimb coordination measures all used dynamic visual stimuli, and thus the correlation between these tests and the SA tests could have been due to the dynamic visual processing required for both types of tests instead of any common time-sharing requirements. Similarly, the temporal processing tests in Experiment 2 also used dynamic visual stimuli, so they could have correlated with SA because of shared visual rather than temporal processing requirements.

In Experiment 3, we attempted to determine to what extent the correlation of both multilimb correlation and temporal processing with SA was due to the visual nature of these tests. For temporal processing, we did this by using auditory tests of temporal processing. If, using these auditory tests, we still find that temporal processing ability predicts SA, then we will have good evidence that this relationship is based on temporal and not visual processing ability.

We could not use the same approach with multilimb coordination as we do not yet have available non-visual tests of multilimb coordination. To determine whether multilimb coordination predicts SA due to its time-sharing or dynamic visual components, we added tests of dynamic visual processing to our battery of cognitive predictors. If we find that dynamic visual processing predicts SA independently of multilimb coordination, this will provide evidence that the multilimb coordination tests predict SA because of their time-sharing rather than visual requirements.

The new tests of dynamic visual processing may also help us understand the relatively low correlation of visual processing with SA in Experiment 2. Experiment 2 found a significant but low correlation between visual search and one SA measure (hazard detection). One possible reason for the low correlation between visual processing and the highly visual SA tasks could be that the visual tests we used required processing of static, unchanging stimuli (e.g., searching for a target in a visual array) while the SA tasks required processing of dynamic, changing visual stimuli. Researchers have found that visual-spatial processing of static and dynamic stimuli are distinguishable abilities (Hunt, Pelligrino, Frick, Farr & Alderton, 1988). Thus, we hypothesized that the dynamic visual processing tests would predict SA independently of static visual processing, and that the dynamic tests would be better predictors than the static ones.

EXPERIMENT 3

In the third experiment, we used the SA tests from Experiments 1 and 2, the performance and recall probes, and added two tests to the SA battery. These were the time-to-passage and scene interpretation tests described earlier. In terms of the SA model in Figure 1, this expanded SA battery included two measures of scene perception and projection and three of local scene comprehension. Also, since the scene interpretation task was done concurrently with the performance and recall tasks, the new battery increased the demands for time-sharing ability.

In terms of cognitive ability tests, we assessed the same four factors as in Experiment 2: working memory, visual, temporal, and perceptual-motor ability. For working memory and perceptual-motor ability, we used tests from the CAM, ASVAB, and Fleishman batteries used in Experiments 1 and 2. However, as mentioned above, we used auditory instead of visual tests to assess temporal ability, and, in addition to tests of static visual processing, added two tests of dynamic visual processing.

Method

Participants. Participants were US Air Force recruits tested during basic training. The 129 participants included 64 males and 65 females ranging in age from 17 to 35 years, with a mean of 20.1 years.

Abilities Tests. In the driving simulator session, participants first completed 32 time-to-passage probes. Then they completed a block of 70 trials, including 30 with only performance probes, 20 with both performance and scene interpretation probes, and 20 with both performance and recall probes. The trials in this block were randomized separately for each participant. Because of this, participants had to be prepared for the performance, recall, and scene-interpretation tasks on each trial.

The multilimb-coordination tests, center the ball and pop the balloons, were described in Experiment 2. The variable representing center-the-ball performance was mean tracking error distance. For pop the balloons, the variable was elapsed time to pop all balloons. The spatial working memory test was similar to the one used in Experiment 2. The participants' ASVAB scores were also obtained.

The first auditory temporal test focused on duration discrimination. On each trial, participants heard an 800 ms tone and either a 620, 692, 764, 836, 908, or 980 ms tone and judged which tone was longer. Tests were scored in terms of percent correct on the four easier discriminations, since participants performed at chance for the two difficult discriminations. In the second test, rate extrapolation, participants heard a series of three or five beeps at one of three rates and then had to press a key when the tenth beep in the series would occur. Performance was scored in terms of the percentage of key presses that were within 0.5 beats of the tenth beat.

In the first static visual search task, spread-out search, participants searched a 3 x 3 array of nine 2-digit numbers, each separated by 9.5 cm, to determine whether any of the numbers

matched a target number. The second task was a computerized version of the AFOQT table reading task. The third was the coding speed test taken from the participants' ASVAB scores. On each item of this test, participants had to determine which of five 4-digit number strings matched a word. To determine which number string matched the word, participants searched a key at the top of the page that consisted of 10 word-number pairs.

In the first dynamic visual test, direction detection, participants saw objects briefly moving across the screen and then, using a response wheel with 16 spokes, indicated the direction in which the object had been moving. In the second test, the road-sign test, participants saw very small letters, numbers, or figures that increased in size. As soon as the object could be identified, the participants pressed a key and then selected the object from a set of distracters.

Procedure. Participants were tested in groups of about 20 who completed the test batteries in a 3.5 hour period. The order of tests was: driving simulator time-to-passage, second driving simulator test, cognitive abilities tests. The order of tests within the cognitive abilities battery was randomized for each participant.

Results and Discussion

Situation Awareness Tests. Table 2 shows how participants performed on the SA tests and the test reliabilities. In this experiment, we used a different measure for hazard detection. Instead of the sensitivity measure used in the previous experiments, hazard detection was estimated by the participants' rate of responding to hazards (hit rate). We ignored false alarm rate and did not calculate sensitivity because the number of catch trials was too low to reliably estimate false alarm rate. Also, three of the items in the scene interpretation test were changed slightly after the first data collection session (N = 28). The scene interpretation percent correct scores for these participants were estimated from the 17 items that remained unchanged.

Table 2. Performance on Situation Awareness Tests for Experiment 3

Measure (units)	Mean	(standard deviation)	Reliability
Time to passage (percent correct)	76	(11)	.63[a]
Composite recall error (car lengths)	0.58	(0.58)	.80[b]
Hazard detection (percent correct)	86	(12)	.80[b]
Blocking-car detection (percent correct)	70	(15)	na
Scene interpretation (percent correct)	51	(15)	.60[a]

[a]Cronbach's alpha.

[b]Corrected even-odd reliability.

The SA scores showed moderate intercorrelations. After switching signs so that positive scores always reflected better performance, the correlations ranged from .10 to .47 with a median of .34. Factor analysis of the SA scores did not reveal the 2-factor structure hypothesized in

Figure 1. Instead we found that a single factor adequately described the SA data ($\chi^2 = 5.51$, $\underline{df} = 5$, $\underline{p} = .36$ for the 1-factor model, 45% of the variance explained).[6] Because of this, we created a *composite SA score*, which was the average of the standardized scores of the five SA tests, with the sign of the standardized composite recall error switched.

Cognitive Correlates of SA. Our goal was to assess the relationships between SA and each of the cognitive, perceptual, and psychomotor tasks, and determine if the abilities measured by these tasks made unique contributions to the prediction of SA in the driving simulator. We began with a factor analysis of the tests intended to measure psychomotor, temporal processing, static visual processing, and dynamic visual processing. To enable the factor analysis program to make fine distinctions among the perceptual and motor tests, we included multiple indicators (two to four) for some of the tests in the analysis. For example, road-sign recognition time scores for letter, figure, and number stimuli were kept separate, and duration discrimination scores for the four easy duration comparison conditions were kept separate.

We excluded certain ASVAB subtests and spatial working memory test from this analysis and used them as indicators of *g/WM* in a separate factor analysis. We estimated *g/WM* as the first unrotated principal axis factor involved in Arithmetic Reasoning, Math Knowledge, Paragraph Comprehension, Word Knowledge, Mechanical Comprehension, and Spatial Working Memory. Note that in this set of tests there were two tests for each of the quantitative, verbal, and spatial domains.

The factor analysis of the perceptual-motor, temporal processing, and visual processing tests was performed in the exploratory mode. To determine the number of factors, we looked for convergence among three criteria: the scree test, the Kaiser-Guttman eigenvalue test, and the maximum likelihood test. Both the scree and the eigenvalue test indicated five factors, but the maximum likelihood test indicated six ($\chi^2 = 117.70$, $\underline{df} = 86$, $\underline{p} = .013$ for 5-factor model). The 5-factor solution (see Table 3) was rotated using an oblique method. We interpreted the factors as multilimb coordination (error/performance time), auditory duration discrimination (accuracy), auditory rate extrapolation (accuracy), dynamic visual processing (time), and visual search (accuracy/rate). Generally, the factors were weakly intercorrelated (median $\underline{r} = .15$), but the two temporal processing factors were correlated .32. The only factor showing substantial correlation with *g/WM* was multilimb coordination ($\underline{r} = -.52$), a negative correlation because motor ability was indexed by errors and performance time. This correlation replicates previous research by Chaiken, Tirre, and Kyllonen (1996), and Ree and Carretta (1994).

Our primary interest was in how *g/WM* and these five perceptual-motor factors combined to predict SA. We computed factor scores using the Anderson-Rubin method which yields uncorrelated factor score estimates. Each of the factors was correlated significantly with the SA composite score (see Table 3), with the highest correlations coming from multilimb coordination (error/performance time) (-.54) and *g/WM* (.45). We then regressed the composite SA score on *g/WM* and the five perceptual/motor factors using the simultaneous inclusion method (Table 4).

[6] If a second factor was extracted (eigenvalue = .94) and an oblique rotation was selected, the factors correlated about .6. For simplicity, we decided to accept the single factor model.

Table 3. Factor Analysis of the Perceptual and Motor Tests

	Multilimb Coordination	Auditory Duration Discrimination	Auditory Rate Extrapolation	Dynamic Visual Processing (Road Sign)	Visual Search
Center-the-Ball (Block 3) Mean error distance	.87				
Center-the-Ball (Block 4) Mean error distance	.86				
Pop-the-Balloons (Block 2) Time	.68				
Pop-the-Balloons (Block 3) Time	.75				
Auditory Duration Discrimination (800 vs. 620 ms) Accuracy		.73			
Auditory Duration Discrimination (800 vs. 692. ms) Accuracy		.66			
Auditory Duration Discrimination (800 vs. 908 ms) Accuracy		.67			
Auditory Duration Discrimination (800 vs. 980 ms) Accuracy		.52			
Auditory Extrapolation (Rate 1) Accuracy			.77		
Auditory Extrapolation (Rate 2) Accuracy			.84		
Auditory Extrapolation (Rate 3) Accuracy			.83		
Road Sign Recognition (Letters) Time				.88	
Road Sign Recognition (Figures) Time				.95	
Road Sign Recognition (Numbers) Time				.92	
Spread-out Search Time					-.38
ASVAB Coding Speed Rate					.37
AFOQT Table Reading Rate					.41
Direction Detection (Even Difficult Blocks) Accuracy	-.49				.50
Direction Detection (Odd Difficult Blocks) Accuracy	-.43				.51
<u>Factor Intercorrelations</u>					
Multilimb Coordination	1.00				
Auditory Duration Discrimination	-.24	1.00			
Auditory Rate Extrapolation	-.09	.32	1.00		
Dynamic Visual Processing	.23	-.19	-.04	1.00	
Visual Search	-.25	.12	-.02	-.07	1.00
<u>Correlations with</u>					
General Cognitive Ability (*g/WM*)	-.52**	.20*	.09	-.12	.24*
Driving Situation Awareness	-.54**	.30**	.18*	-.29**	.18*

<u>Note</u>. $\underline{N} = 111$. Loadings < .30 omitted. *$\underline{p} < .05$, **$\underline{p} < .01$. $\underline{r}$(*g/WM*, SA) = .445.

The six predictors together accounted for 52.9% of the variance in the SA composite score ($\underline{R}$ = .73, $\underline{R}_{adj}$ = .71). All predictors except *g/WM* made significant unique contributions to the equation. The primary contest for criterion variance was between *g/WM* and multilimb coordination. When these two variables were used to predict the SA composite score in a two-

variable equation, multilimb coordination contributed .13 to the proportion of variance explained, and *g/WM* contributed .04. The common contribution by *g/WM* and multilimb coordination was .16.

Table 4. Regression Analysis: Prediction of Driving Situation Awareness Composite

Variable	Beta	r	Part r	Partial r	t
General Cognitive Ability (*g/WM*)	.02	.45	.02	.02	.23
Multilimb Coordination (Error/Performance Time)	-.53	-.54	-.44	-.54	-6.38**
Auditory Duration Discrimination	.30	.30	.29	.39	4.17**
Auditory Rate Extrapolation	.18	.18	.18	.25	2.59*
Dynamic Visual Processing	-.29	-.29	-.28	-.38	-4.12**
Visual Search	.17	.18	.17	.24	2.42*

Note. $N = 106$. $R = .73$, $R_{adj} = .71$. $F(6, 99) = 18.56$, $p < .0001$. $*p < .05$, $**p < .001$

The results of the regression analysis are consistent with the idea that SA is a product of several cognitive and perceptual processes--time-sharing (indicated through multilimb coordination), temporal processing, and visual processing. Experiment 3 also provided information about the reasons behind the correlations of temporal processing and multilimb coordination with SA. The fact that both auditory temporal processing tests made significant contributions to predicting SA suggests that temporal processing is an important part of SA and that the correlation between temporal processing and SA in Experiment 2 was not solely due to shared visual processing.

Concerning multilimb coordination, we hypothesized earlier that it is strongly related to SA either because both these tasks require time-sharing capacity or because both tasks require dynamic visual processing. The evidence from Experiment 3 appears to favor the first hypothesis because multilimb coordination and both the dynamic and static visual processing factors all made unique contributions to the explained variance in SA. It should be noted that we do not possess direct evidence to support the interpretation of multilimb coordination as reflecting time-sharing capacity. However, Baddeley (1993) described working memory as involving the management of cognitive resources or time-sharing, and the substantial correlation between *g/WM* and multilimb coordination is what we would expect if multilimb coordination involved time-sharing as well.

Finally, Experiment 3 provided mixed evidence on the question of whether dynamic visual processing predicts SA independently of static visual processing. One dynamic visual test, the road-sign test, was grouped by the factor analysis in a separate factor from the static visual tests and was found by the regression analysis to predict SA independently of the static tests.

The other dynamic visual test, direction detection, was grouped by the factor analysis with both dynamic tests, *viz.*, pop-the-balloons and center-the-ball, and static visual search. This suggests that direction detection involves two visual components: (a) a dynamic visual component distinct from the component involved in the road sign test, but required by visual tracking tasks such as center-the-ball and pop-the-balloons, and (b) a static visual component similar to that required in visual search tasks. If we accept this interpretation of variance on the direction detection test, we can conclude that both dynamic and static visual processes are involved in SA.

We had also hypothesized that dynamic visual tests would be better predictors of SA than static visual tests. However, the correlation between SA and the dynamic visual processing (road-sign) factor (r = .29) was not significantly higher than that between SA and the visual search factor that included the static tests (r = .18, $t(108) = 0.88$, $p < .20$).

GENERAL DISCUSSION

In these studies we measured SA using a low-fidelity driving simulator in which participants were required to recall car locations and to identify and avoid hazards. Cognitive abilities were measured using test batteries that focused on working-memory capacity and processing speed in Experiment 1, and on working memory and visual, temporal, and perceptual-motor processing in Experiments 2 and 3.

Only working memory was found to be strongly related to SA in Experiment 1, but statistical power was low due to a small sample size and relationships with other predictors may have been obscured. In Experiment 2, we employed a larger sample and found that working memory was strongly predictive of both SA dependent variables employed. We also found that multilimb coordination added to working memory in predicting one SA dependent variable (blocking car detection). Multilimb coordination, interpreted as an indicator for time-sharing ability, and working memory each made unique contributions to the explained variance in SA. However, their shared or common contribution was also substantial, indicating that there is a high degree of commonality between these variables as Baddeley (1993) suggested.

On the surface, the results of Experiment 3 might appear to contradict the tentative conclusion that working memory plays a substantial role in SA. In Experiment 3 we found that *g/WM* failed to make a significant unique contribution to explanation of SA variance. Unfortunately, we did not measure *g/WM* in quite the same manner as in Experiments 1 and 2, relying instead on the *g* component found among selected ASVAB tests along with a spatial working memory measure. This decision probably weakened our measure of working memory to some degree. Multilimb coordination overlapped substantially with *g/WM* and in regression equations may have "stolen" some of the variance *g/WM* would have explained on its own. The shared contribution of these two variables was larger than either variable's unique contribution.

Working memory, as it is the central bottleneck in controlled cognitive processing, would be expected to be involved in conscious processing leading to SA; but as noted by other researchers, SA might have an automatic component as well, that might not be limited by working memory

(Kennedy & Ordy, 1995; Orasanu, 1996; Orasnu & Fischer, 1997). This automatic component of SA merits further investigation.

Visual processing was found to predict SA ability in Experiments 2 and 3. The evidence appears to suggest that both dynamic (e.g., road sign recognition) and static (e.g., visual search) processing factors are involved in SA. Our understanding of the visual processing factors operating in the dynamic driving environment is admittedly incomplete, and we hope to refine our measurement of these abilities in future research. Temporal processing predicted SA ability in Experiments 2 and 3, regardless of whether temporal ability was measured by visual or auditory tests. Experiment 3 demonstrated that temporal processing might be multicomponential itself. Two temporal factors were found in the factor analysis and each contributed significantly to the prediction of SA.

Our findings that SA ability is correlated with *g*, working memory ability, and perceptual - motor ability fit with the findings of Carretta et al. (1996) and Endsley and Bolstad (1994), and also with our findings that the real-time defensive systems operator task is predicted by g. Our finding that SA ability is correlated with static visual processing fits with O'Hare's (1997) finding that SA is predicted by visual recognition ability. In addition, we also found that SA ability is correlated with dynamic visual processing ability and temporal processing ability.

As the next steps in our research program, we propose to pursue two streams simultaneously. One involves refining our measures of the cognitive correlates of SA and exploring relations among these. As suggested earlier, we believe that time-sharing, dynamic visual processing, and temporal processing abilities each need further research. The second line of research involves refining our measurement of SA. A future version of the driving simulator will introduce interactive driving and multitasking to increase realism and generalizability to automobile driving. Research participants will be able to steer, brake, accelerate, and navigate their simulated cars through a simulated city environment. The multitasking requirement will be increased through introduction of navigation and communication tasks. Imagine the workload of a sales representative trying to find a location in an unfamiliar neighborhood who must respond to calls from the office on his cellular phone. We can simulate these conditions and measure the effect on SA.

An important part of refining our SA measures will involve validating that these measures predict performance of real-time tasks under realistic conditions. This validation is necessary to support the claims made earlier that maintaining SA is a crucial part of real-time task performance. In our case, this would require predicting performance in on-the-road driving. As mentioned earlier, Gopher et al. (1992) and Hart and Battiste (1992) found that practice on Space Fortress benefited subsequent aircraft piloting performance. Similarly, performance on the WOMBAT test of SA differentiated between elite pilots and nonpilots (O'Hare, 1997).

SA is a concept that first emerged in the aviation community, and recently, human factors researchers have extended this concept to many domains of human performance. We initially chose driving as a simpler and more common analogue to flying, thinking that if we could first understand SA in driving, we would have a good foundation for understanding SA in flying. Our

plans are to examine correlations between measures of SA in the driving simulator and measures of SA in other simulated environments, e.g., air combat and air traffic control or weapons direction. Confirmatory factor analysis will be used to test alternative structural models of the cognitive processes involved in SA. Also in this research, we will determine if similar predictors (cognitive correlates) emerge for SA in different environments.

Future applied research should investigate the utility of SA measures and SA correlates for personnel selection, training, and human-system interface design. An example of a system that is likely to benefit from this research is the uninhabited air vehicle (UAV). UAVs represent an emerging technology in which ground-control station crew concepts, training content and methods, and human-system interfaces for the various crewmembers all present significant challenges to system effectiveness. Cognitive task analyses conducted with air vehicle operators (pilots) and payload operators for the Predator (RQ-1) have identified situation awareness as a major concern of operators that merits research attention (Hall & Tirre, 1998).

Gopher, D., Weil, M., Bareket, T. (1992). The transfer of skill from a computer game trainer to actual flight (pp. 1285-1290). In Symposium: Training: Development of Skill Trainers Based on Computer Games: Prospects and Issues. Proceedings of the Human Factors Society 36th Annual Meeting. Santa Monica, CA: Human Factors Society.

Grier, J. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. Psychological Bulletin, 75, 424-429.

Gugerty, L. (1997). Situation awareness during driving: Explicit and implicit knowledge in dynamic spatial memory. Journal of Experimental Psychology: Applied, 3, 1-26.

Gugerty, L. & Tirre, W. (1996). Situation awareness: A validation study and investigation of individual differences. In Proceedings of the 40th Annual Meeting of the Human Factors Society (pp. 564-568). Santa Monica, CA: Human Factors and Ergonomics Society.

Hall, E. & Tirre, W.C. (1998). USAF air vehicle operator training requirements study. Brooks Air Force Base, TX: (AFRL-HE-BR-SR-1998-0001). Brooks Air Force Base, TX: Air Force Research Laboratory, Human Effectiveness Directorate.

Hart, S. & Battiste, V. (1992). Field test of video game trainer. In Symposium: Training: Development of Skill Trainers Based on Computer Games: Prospects and Issues (pp. 1291-1295). Proceedings of the Human Factors Society 36th Annual Meeting. Santa Monica, CA: Human Factors Society.

Hartel, C., Smith, K. & Prince, C. (1991). Defining aircrew coordination: Searching for mishaps with meaning. Sixth International Symposium on Aviation Psychology. Columbus, OH: Department of Aviation, Ohio State University.

Hopkin, D. (1993, February). Situation awareness in air traffic control. Paper presented at the Situation Awareness in Complex Systems Conference. Orlando, FL: Center for Applied Human Factors in Aviation.

Hunt, E., Pelligrino, J., Frick, R., Farr, S. & Alderton, D. (1988). The ability to reason about movement in the visual field. Intelligence, 12, 77-100.

Joslyn, S. & Hunt, E. (1998). Evaluating individual differences in response to time-pressure situations. Journal of Experimental Psychology: Applied, 4(1), 16-43.

Kaempf, G., Klein, G., Thordsen, M. & Wolf, S. (1996). Decision making in complex naval command-and-control environments. Human Factors, 38, 220-231.

Kennedy, R. & Ordy, J. M. (1995). Situation awareness: A cognitive neuroscience model based on specific neurobehavioral mechanisms. In D. Garland & M. Endsley (Eds.), Experimental Analysis and Measurement of Situation Awareness (pp. 49-56). Daytona Beach, FL: Embry-Riddle Aeronautical University Press.

Klein, G. (1993). A recognition-primed decision (RPD) model of rapid decision making. In G. Klein, J. Orasanu, R. Calderwood & C. Zsambok (Eds.), Decision making in action: Models and methods (pp. 138-147). Norwood, NJ: Ablex.

Kyllonen, P.C. (1993). Aptitude testing based on information processing: A test of the four-sources model. Journal of General Psychology, 120, 375-405.

Kyllonen, P.C. (1994). CAM: A theoretical framework for cognitive abilities measurement. In D. Detterman (Ed.), Current topics in human intelligence: Volume IV, Theories of intelligence (pp. 307-359). Norwood, NJ: Ablex.

Kyllonen, P. C. & Christal, R. E. (1990). Reasoning ability is (little more than) working memory capacity? Intelligence, 14, 389-433.

O'Hare, D. (1997). Cognitive ability determinants of elite pilot performance. Human Factors, 39, 54-552.

Orasanu, J. (1996). Finding decisions in natural environments: A view from the cockpit. In C. Zsambok & G. Klein (Eds.), Naturalistic decision making. Mahwah, NJ: Erlbaum.

Orasanu, J. & Fischer, U. (1997). Finding decisions in natural environments: Towards a theory of situated decision making. In C. Zsambok & G. Klein (Eds.), Naturalistic decision making. (pp. 343-357) Mahwah, NJ: Erlbaum.

Rabbitt, P., Banerji, N., & Szymanski, A. (1989). Space Fortress as an IQ test? Predictions of learning and of practice performance in a complex interactive video-game. Acta Psychologica, 71, 243-257.

Ree, M. J., & Carretta, T. R. (1994). The correlation of general cognitive ability and psychomotor tracking tests. International Journal of Selection and Assessment, 2, 209-216.

Roscoe, S.N. (1993). Predicting and enhancing flight deck performance. In R. Telfer & P. Moore (Eds.), Aviation training: Pilot, instructor and organisation (pp. 195-208). Aldershot, England: Avebury.

Roscoe, S.N. & Corl, L. (1987). Wondrous original method for basic airmanship testing. In R.S. Jensen (Ed.), Proceedings of the 4th International Symposium on Aviation Psychology (pp. 493-499). Columbus, OH: Ohio State University, Department of Aviation.

Shinar, D. (1993). Traffic safety and individual differences in drivers' attention and information processing capacity. Alcohol, Drugs and Driving, 9, 219-237.

Spearman, C.E. (1904). 'General intelligence' objectively determined and measured. American Journal of Psychology, 15, 201-293.

Spearman, C.E. (1923). The nature of 'intelligence' and the principles of cognition. 2nd Edition, London: MacMillan. (Reprinted in 1973 by Arno Press).

Tirre, W. C. (1997). Steps toward an improved pilot selection battery. In R. Dillon (Ed.), Handbook on Testing (pp. 220-255). Westport, CT: Greenwood Press.

Tirre, W. C., & Pena, M. C. (1993). Components of quantitative reasoning: General and group factors. Intelligence, 17, 501-522.

Watson, C. & Nichols, T. (1976). Detectability of auditory signals presented without defined observation intervals. Journal of the Acoustical Society of America, 59, 655-667.